OIPE

#2

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/804,626

DATE: 03/30/2001
TIME: 15:02:53

RECEIVED APR 11 2001 TECHCENTER 1600/2900

ENTERED

Input Set : **A:\PTO.txt**
Output Set: **N:\CRF3\03302001\I804626.raw**

```
   3  <110> APPLICANT: Lobel, Leslie
   4        Lustbader, Joyce
   6  <120> TITLE OF INVENTION: EXPRESSION OF PROPERLY FOLDED AND SOLUBLE EXTRACELLULAR
DOMAIN OF A
   7        GONADOTROPIN RECEPTOR
   9  <130> FILE REFERENCE: 0575/62259/JPW/SHS
C--> 11 <140> CURRENT APPLICATION NUMBER: US/09/804,626
C--> 11 <141> CURRENT FILING DATE: 2001-03-09
     11 <160> NUMBER OF SEQ ID NOS: 8
     13 <170> SOFTWARE: PatentIn version 3.0
     15 <210> SEQ ID NO: 1
     16 <211> LENGTH: 1557
     17 <212> TYPE: DNA
     18 <213> ORGANISM: HOMO SAPIENS
     20 <400> SEQUENCE: 1
     21 atgagcgata aaattattca cctgactgac gacagttttg acacggatgt actcaaagcg      60
     23 gacggggcga tcctcgtcga tttctgggca gagtggtgcg gtccgtgcaa aatgatcgcc     120
     25 ccgattctgg atgaaatcgc tgacgaatat cagggcaaac tgaccgttgc aaaactgaac     180
     27 atcgatcaaa accctggcac tgcgccgaaa tatggcatcc gtggtatccc gactctgctg     240
     29 ctgttcaaaa acggtgaagt ggcggcaacc aaagtgggtg cactgtctaa aggtcagttg     300
     31 aaagagttcc tcgacgctaa cctggccggt tctggttctg gccatatgca ccatcatcat     360
     33 catcattctt ctggtctggt gccacgcggt tctggtatga agaaaccgc tgctgctaaa     420
     35 ttcgaacgcc agcacatgga cagcccagat ctgggtaccg acgacgacga caaggccatg     480
     37 gctgatatcg gatccgaatt ccgagcgctg cgcgaggcgc tctgccctga ccctgcaac     540
     39 tgcgtgcccg acggcgccct cgcctgcccc ggccccacgg ccggtctcac tcgactatca     600
     41 cttgcctacc tccctgtcaa agtgatccca tctcaagctt tcagaggact taatgaggtc     660
     43 ataaaaattg aaatctctca gattgattcc ctggaaagga tagaagctaa tgcctttgac     720
     45 aacctcctca atttgtctga atactgatc cagaacacca aaaatctgag atacattgag     780
     47 cccggagcat ttataaatct tccccgatta aaatacttga gcatctgtaa cacaggcatc     840
     49 agaaagtttc cagatgttac gaaggtcttc tcctctgaat caaatttcat tctggaaatt     900
     51 tgtgataact acacataac caccatacca ggaaatgctt ttcaagggat gaataatgaa     960
     53 tctgtaacac tcaaactata tggaaatgga tttgaagaag tacaaagtca tgcattcaat    1020
     55 gggacgacac tgacttcact ggagctaaag gaaaacgtac atctggagaa gatgcacaat    1080
     57 ggagccttcc gtgggggccac agggccgaaa accttggata tttcttccac caaattgcag    1140
     59 gccctgccga gctatggcct agagtccatt cagaggctaa ttgccacgtc atcctattct    1200
     61 ctaaaaaaat tgccatcaag agaaacattt gtcaatctcc tggaggccac gttgacttac    1260
     63 cccagccact gctgtgcttt tagaaacttg ccaacaaaag aacagaattt ttcacattcc    1320
     65 atttctgaaa acttttccaa acaatgtgaa agcacagtaa ggaaagtgaa taacaaaaca    1380
     67 ctttattctt ccatgcttgc tgagagtgaa ctgagtggct gggactatga atatggtttc    1440
     69 tgcttaccca agacacccccg atgtgctcct gaaccagatg cttttaatcc ctgtgaagat    1500
     71 attatgggcg tcgacaagct tgcggccgca ctcgagcacc accaccacca ccactga      1557
     74 <210> SEQ ID NO: 2
     75 <211> LENGTH: 518
     76 <212> TYPE: PRT
     77 <213> ORGANISM: Homo Sapiens
     79 <400> SEQUENCE: 2
     81 Met Ser Asp Lys Ile Ile His Leu Thr Asp Asp Ser Phe Asp Thr Asp
     82 1               5                   10                  15
```

RAW SEQUENCE LISTING   DATE: 03/30/2001
PATENT APPLICATION: US/09/804,626   TIME: 15:02:53

Input Set : **A:\PTO.txt**
Output Set: **N:\CRF3\03302001\I804626.raw**

```
84 Val Leu Lys Ala Asp Gly Ala Ile Leu Val Asp Phe Trp Ala Glu Trp
85             20              25              30
87 Cys Gly Pro Cys Lys Met Ile Ala Pro Ile Leu Asp Glu Ile Ala Asp
88             35              40              45
90 Glu Tyr Gln Gly Lys Leu Thr Val Ala Lys Leu Asn Ile Asp Gln Asn
91     50              55              60
93 Pro Gly Thr Ala Pro Lys Tyr Gly Ile Arg Gly Ile Pro Thr Leu Leu
94 65              70              75              80
96 Leu Phe Lys Asn Gly Glu Val Ala Ala Thr Lys Val Gly Ala Leu Ser
97             85              90              95
99 Lys Gly Gln Leu Lys Glu Phe Leu Asp Ala Asn Leu Ala Gly Ser Gly
100            100             105             110
102 Ser Gly His Met His His His His His His Ser Ser Gly Leu Val Pro
103            115             120             125
105 Arg Gly Ser Gly Met Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln
106     130             135             140
108 His Met Asp Ser Pro Asp Leu Gly Thr Asp Asp Asp Asp Lys Ala Met
109 145            150             155             160
111 Ala Asp Ile Gly Ser Glu Phe Arg Ala Leu Arg Glu Ala Leu Cys Pro
112            165             170             175
114 Glu Pro Cys Asn Cys Val Pro Asp Gly Ala Leu Arg Cys Pro Gly Pro
115            180             185             190
117 Thr Ala Gly Leu Thr Arg Leu Ser Leu Ala Tyr Leu Pro Val Lys Val
118     195             200             205
120 Ile Pro Ser Gln Ala Phe Arg Gly Leu Asn Glu Val Ile Lys Ile Glu
121     210             215             220
123 Ile Ser Gln Ile Asp Ser Leu Glu Arg Ile Glu Ala Asn Ala Phe Asp
124 225            230             235             240
126 Asn Leu Leu Asn Leu Ser Glu Ile Leu Ile Gln Asn Thr Lys Asn Leu
127            245             250             255
129 Arg Tyr Ile Glu Pro Gly Ala Phe Ile Asn Leu Pro Arg Leu Lys Tyr
130            260             265             270
132 Leu Ser Ile Cys Asn Thr Gly Ile Arg Lys Phe Pro Asp Val Thr Lys
133     275             280             285
135 Val Phe Ser Ser Glu Ser Asn Phe Ile Leu Glu Ile Cys Asp Asn Leu
136     290             295             300
138 His Ile Thr Thr Ile Pro Gly Asn Ala Phe Gln Gly Met Asn Asn Glu
139 305            310             315             320
141 Ser Val Thr Leu Lys Leu Tyr Gly Asn Gly Phe Glu Glu Val Gln Ser
142            325             330             335
144 His Ala Phe Asn Gly Thr Thr Leu Thr Ser Leu Glu Leu Lys Glu Asn
145            340             345             350
147 Val His Leu Glu Lys Met His Asn Gly Ala Phe Arg Gly Ala Thr Gly
148     355             360             365
150 Pro Lys Thr Leu Asp Ile Ser Ser Thr Lys Leu Gln Ala Leu Pro Ser
151     370             375             380
153 Tyr Gly Leu Glu Ser Ile Gln Arg Leu Ile Ala Thr Ser Ser Tyr Ser
154 385            390             395             400
156 Leu Lys Lys Leu Pro Ser Arg Glu Thr Phe Val Asn Leu Leu Glu Ala
```

```
157                   405                    410                     415
159 Thr Leu Thr Tyr Pro Ser His Cys Cys Ala Phe Arg Asn Leu Pro Thr
160               420                    425                    430
162 Lys Glu Gln Asn Phe Ser His Ser Ile Ser Glu Asn Phe Ser Lys Gln
163           435                    440                    445
165 Cys Glu Ser Thr Val Arg Lys Val Asn Asn Lys Thr Leu Tyr Ser Ser
166       450                    455                    460
168 Met Leu Ala Glu Ser Glu Leu Ser Gly Trp Asp Tyr Glu Tyr Gly Phe
169 465               470                    475                    480
171 Cys Leu Pro Lys Thr Pro Arg Cys Ala Pro Glu Pro Asp Ala Phe Asn
172           485                    490                    495
174 Pro Cys Glu Asp Ile Met Gly Val Asp Lys Leu Ala Ala Ala Leu Glu
175           500                    505                    510
177 His His His His His His
178           515
180 <210> SEQ ID NO: 3
181 <211> LENGTH: 1551
182 <212> TYPE: DNA
183 <213> ORGANISM: Homo Sapiens

185 <400> SEQUENCE: 3
186 atgagcgata aaattattca cctgactgac gacagttttg acacggatgt actcaaagcg      60
188 gacggggcga tcctcgtcga tttctgggca gagtggtgcg gtccgtgcaa aatgatcgcc     120
190 ccgattctgg atgaaatcgc tgacgaatat cagggcaaac tgaccgttgc aaaactgaac     180
192 atcgatcaaa accctggcac tgcgccgaaa tatggcatcc gtggtatccc gactctgctg     240
194 ctgttcaaaa acggtgaagt ggcggcaacc aaagtgggtg cactgtctaa aggtcagttg     300
196 aaagagttcc tcgacgctaa cctggccggt tctggttctg gccatatgca ccatcatcat     360
198 catcattctt ctggtctggt gccacgcggt tctggtatga agaaaccgc tgctgctaaa      420
200 ttcgaacgcc agcacatgga cagcccagat ctgggtaccg acgacgacga caaggccatg     480
202 gctgatatcg gatccgaatt cagggttttt ctctgccaag agagcaaggt gacagagatt     540
204 ccttctgacc tcccgaggaa tgccattgaa ctgaggtttg tcctcaccaa gcttcgagtc     600
206 atccaaaaag gtgcattttc aggatttggg gacctggaga aaatagagat ctctcagaat     660
208 gatgtcttgg aggtgataga ggcagatgtg ttctccaacc ttcccaaatt acatgaaatt     720
210 agaattgaaa aggccaacaa cctgctctac atcacccctg aggccttcca gaaccttccc     780
212 aaccttcaat atctgttaat atccaacaca ggtattaagc accttccaga tgttcacaag     840
214 attcattctc tccaaaaggt tttacttgac attcaagata acataaacat ccacacaatt     900
216 gaaagaaatt ctttcgtggg gctgagcttt gaaagtgtga ttctatggct gaataagaat     960
218 gggattcaag aaatacacaa ctgtgcattc aatggaaccc aactagatgc agtgaatcta    1020
220 agcgataata ataatttaga agaattgcct aatgatgttt ccacggagc ctctggacca     1080
222 gtcattctag atatttcaag aacaaggatc cattccctgc ctagctatgg cttagaaaat    1140
224 cttaagaagc tgagggccag gtcgacttac aacttaaaaa agctgcctac tctggaaaag    1200
226 cttgtcgccc tcatggaagc cagcctcacc tatcccagcc attgctgtgc ctttgcaaac    1260
228 tggagacggc aaatctctga gcttcatcca atttgcaaca aatctatttt aaggcaagaa    1320
230 gttgattata tgactcaggc tagggggtcag agatcctctc tggcagaaga caatgagtcc    1380
232 agctacagca gaggatttga catgacgtac actgagtttg actatgactt atgcaatgaa    1440
234 gtggttgacg tgacctgctc ccctaagcca gatgcattca acccatgtga agatatcatg    1500
236 ggggtcgaca agcttgcggc cgcactcgag caccaccacc accaccactg a              1551
239 <210> SEQ ID NO: 4
240 <211> LENGTH: 516
241 <212> TYPE: PRT
```

```
242 <213> ORGANISM: Homo Sapiens
244 <400> SEQUENCE: 4
246 Met Ser Asp Lys Ile Ile His Leu Thr Asp Asp Ser Phe Asp Thr Asp
247 1               5                   10                  15
249 Val Leu Lys Ala Asp Gly Ala Ile Leu Val Asp Phe Trp Ala Glu Trp
250             20                  25                  30
252 Cys Gly Pro Cys Lys Met Ile Ala Pro Ile Leu Asp Glu Ile Ala Asp
253         35                  40                  45
255 Glu Tyr Gln Gly Lys Leu Thr Val Ala Lys Leu Asn Ile Asp Gln Asn
256     50                  55                  60
258 Pro Gly Thr Ala Pro Lys Tyr Gly Ile Arg Gly Ile Pro Thr Leu Leu
259 65                  70                  75                  80
261 Leu Phe Lys Asn Gly Glu Val Ala Ala Thr Lys Val Gly Ala Leu Ser
262             85                  90                  95
264 Lys Gly Gln Leu Lys Glu Phe Leu Asp Ala Asn Leu Ala Gly Ser Gly
265             100                 105                 110
267 Ser Gly His Met His His His His His His Ser Ser Gly Leu Val Pro
268         115                 120                 125
270 Arg Gly Ser Gly Met Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln
271     130                 135                 140
273 His Met Asp Ser Pro Asp Leu Gly Thr Asp Asp Asp Asp Lys Ala Met
274 145                 150                 155                 160
276 Ala Asp Ile Gly Ser Glu Phe Arg Val Phe Leu Cys Gln Glu Ser Lys
277             165                 170                 175
279 Val Thr Glu Ile Pro Ser Asp Leu Pro Arg Asn Ala Ile Glu Leu Arg
280             180                 185                 190
282 Phe Val Leu Thr Lys Leu Arg Val Ile Gln Lys Gly Ala Phe Ser Gly
283         195                 200                 205
285 Phe Gly Asp Leu Glu Lys Ile Glu Ile Ser Gln Asn Asp Val Leu Glu
286     210                 215                 220
288 Val Ile Glu Ala Asp Val Phe Ser Asn Leu Pro Lys Leu His Glu Ile
289 225                 230                 235                 240
291 Arg Ile Glu Lys Ala Asn Asn Leu Leu Tyr Ile Thr Pro Glu Ala Phe
292             245                 250                 255
294 Gln Asn Leu Pro Asn Leu Gln Tyr Leu Leu Ile Ser Asn Thr Gly Ile
295             260                 265                 270
297 Lys His Leu Pro Asp Val His Lys Ile His Ser Leu Gln Lys Val Leu
298     275                 280                 285
300 Leu Asp Ile Gln Asp Asn Ile Asn Ile His Thr Ile Glu Arg Asn Ser
301     290                 295                 300
303 Phe Val Gly Leu Ser Phe Glu Ser Val Ile Leu Trp Leu Asn Lys Asn
304 305                 310                 315                 320
306 Gly Ile Gln Glu Ile His Asn Cys Ala Phe Asn Gly Thr Gln Leu Asp
307             325                 330                 335
309 Ala Val Asn Leu Ser Asp Asn Asn Leu Glu Glu Leu Pro Asn Asp
310             340                 345                 350
312 Val Phe His Gly Ala Ser Gly Pro Val Ile Leu Asp Ile Ser Arg Thr
313         355                 360                 365
315 Arg Ile His Ser Leu Pro Ser Tyr Gly Leu Glu Asn Leu Lys Lys Leu
```

RAW SEQUENCE LISTING   DATE: 03/30/2001
PATENT APPLICATION: US/09/804,626   TIME: 15:02:53

Input Set : A:\PTO.txt
Output Set: N:\CRF3\03302001\I804626.raw

```
316      370                     375                     380
318 Arg Ala Arg Ser Thr Tyr Asn Leu Lys Lys Leu Pro Thr Leu Glu Lys
319 385                     390                     395                     400
321 Leu Val Ala Leu Met Glu Ala Ser Leu Thr Tyr Pro Ser His Cys Cys
322                 405                     410                     415
324 Ala Phe Ala Asn Trp Arg Arg Gln Ile Ser Glu Leu His Pro Ile Cys
325                 420                     425                     430
327 Asn Lys Ser Ile Leu Arg Gln Glu Val Asp Tyr Met Thr Gln Ala Arg
328         435                     440                     445
330 Gly Gln Arg Ser Ser Leu Ala Glu Asp Asn Glu Ser Ser Tyr Ser Arg
331     450                     455                     460
333 Gly Phe Asp Met Thr Tyr Thr Glu Phe Asp Tyr Asp Leu Cys Asn Glu
334 465                     470                     475                     480
336 Val Val Asp Val Thr Cys Ser Pro Lys Pro Asp Ala Phe Asn Pro Cys
337                 485                     490                     495
339 Glu Asp Ile Met Gly Val Asp Lys Leu Ala Ala Ala Leu Glu His His
340                 500                     505                     510
342 His His His His
343         515
345 <210> SEQ ID NO: 5
346 <211> LENGTH: 2100
347 <212> TYPE: DNA
348 <213> ORGANISM: Homo Sapiens
350 <400> SEQUENCE: 5
351 atgaagcagc ggttctcggc gctgcagctg ctgaagctgc tgctgctgct gcagccgccg      60
353 ctgccacgag cgctgcgcga ggcgctctgc cctgagccct gcaactgcgt gcccgacggc     120
355 gccctgcgct gccccggccc cacggccggt ctcactcgac tatcacttgc ctacctccct     180
357 gtcaaagtga tcccatctca agctttcaga ggacttaatg aggtcataaa aattgaaatc     240
359 tctcagattg attccctgga aaggatagaa gctaatgcct ttgacaacct cctcaatttg     300
361 tctgaaatac tgatccagaa caccaaaaat ctgagataca ttgagcccgg agcatttata     360
363 aatcttcccc gattaaaata cttgagcatc tgtaacacag gcatcagaaa gtttccagat     420
365 gttacgaagg tcttctcctc tgaatcaaat ttcattctgg aaatttgtga taacttacac     480
367 ataaccacca taccaggaaa tgcttttcaa gggatgaata atgaatctgt aacactcaaa     540
369 ctatatggaa atggatttga agaagtacaa agtcatgcat tcaatgggac gacactgact     600
371 tcactggagc taaaggaaaa cgtacatctg gagaagatgc acaatggagc cttccgtggg     660
373 gccacagggc cgaaaacctt ggatatttct tccaccaaat gcaggccct gccgagctat     720
375 ggcctagagt ccattcagag gctaattgcc acgtcatcct attctctaaa aaaattgcca     780
377 tcaagagaaa catttgtcaa tctcctggag gccacgttga cttaccccag ccactgctgt     840
379 gcttttagaa acttgccaac aaaagaacag aattttttcac attccatttc tgaaaacttt     900
381 tccaaacaat gtgaaagcac agtaaggaaa gtgaataaca aaacacttta ttcttccatg     960
383 cttgctgaga gtgaactgag tggctgggac tatgaatatg gtttctgctt acccaagaca    1020
385 ccccgatgtg ctcctgaacc agatgctttt aatccctgtg aagatattat gggctatgac    1080
387 ttccttaggg tcctgatttg gctgattaat attctagcca tcatgggaaa catgactgtt    1140
389 cttttgttc tcctgacaag tcgttacaaa cttacagtgc ctcgtttttct catgtgcaat    1200
391 ctctcctttg cagacttttg catgggctc tatctgctgc tcatagcctc agttgattcc    1260
393 caaaccaagg gccagtacta taaccatgcc atagactggc agacagggag tgggtgcagc    1320
395 actgctggct ttttcactgt attcgcaagt gaactttctg tctacaccct caccgtcatc    1380
397 actctagaaa gatggcacac catcacctat gctattcacc tggaccaaaa gctgcgatta    1440
399 agacatgcca ttctgattat gcttggagga tggctctttt cttctctaat tgctatgttg    1500
```

VERIFICATION SUMMARY   DATE: 03/30/2001
PATENT APPLICATION: **US/09/804,626**   TIME: 15:02:54

Input Set : **A:\PTO.txt**
Output Set: **N:\CRF3\03302001\I804626.raw**

L:11 M:270 C: Current Application Number differs, Replaced Current Application No
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date